RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 09/508,658C
Source: IFW16
Date Processed by STIC: 10/3/05

ENTERED

IFW16

RAW SEQUENCE LISTING DATE: 10/03/2005
PATENT APPLICATION: **US/09/508,658C** TIME: 13:29:19

Input Set : **A:\u012653-9.txt**
Output Set: **N:\CRF4\10032005\I508658C.raw**

```
 3 <110> APPLICANT: KROHN, Kai
 4       HEINO, Maarit
 5       PETERSON, Part
 6       SCOTT, Hamish
 7       ANTONARAKIS, Stylianos
 8       LALIOTI, Maria D.
 9       SHIMIZU, Nobuyoshi D.
10       KUDOH, Jun D.
12 <120> TITLE OF INVENTION: NOVEL GENE DEFECTIVE IN APECED AND ITS USE
14 <130> FILE REFERENCE: u 012653-9
16 <140> CURRENT APPLICATION NUMBER: 09/508,658C
17 <141> CURRENT FILING DATE: 2000-11-03
19 <160> NUMBER OF SEQ ID NOS: 39
21 <170> SOFTWARE: PatentIn version 3.2
23 <210> SEQ ID NO: 1
24 <211> LENGTH: 2036
25 <212> TYPE: DNA
26 <213> ORGANISM: HOMO SAPIENS
29 <220> FEATURE:
30 <221> NAME/KEY: CDS
31 <222> LOCATION: (137)..(1771)
32 <223> OTHER INFORMATION: /product="AIR-1"
34 <400> SEQUENCE: 1
35 agaccgggga gacgggcggg cgcacagccg gcgcggaggc cccacagccc cgccgggacc     60
37 cgaggccaag cgaggggctg ccagtgtccc gggacccacc gcgtccgccc cagccccggg    120
39 tccccgcgcc cacccc atg gcg acg gac gcg gcg cta cgc cgg ctt ctg agg    172
40                    Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg
41                    1               5                   10
43 ctg cac cgc acg gag atc gcg gtg gcc gtg gac agc gcc ttc cca ctg      220
44 Leu His Arg Thr Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu
45         15                  20                  25
47 ctg cac gcg ctg gct gac cac gac gtg gtc ccc gag gac aag ttt cag      268
48 Leu His Ala Leu Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln
49     30                  35                  40
51 gag acg ctt cat ctg aag gaa aag gag ggc tgc ccc cag gcc ttc cac      316
52 Glu Thr Leu His Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His
53 45                  50                  55                  60
55 gcc ctc ctg tcc tgg ctg ctg acc cag gac tcc aca gcc atc ctg gac      364
56 Ala Leu Leu Ser Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp
57                 65                  70                  75
59 ttc tgg agg gtg ctg ttc aag gac tac aac ctg gag cgc tat ggc cgg      412
60 Phe Trp Arg Val Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg
61             80                  85                  90
```

RAW SEQUENCE LISTING DATE: 10/03/2005
PATENT APPLICATION: **US/09/508,658C** TIME: 13:29:19

Input Set : **A:\u012653-9.txt**
Output Set: **N:\CRF4\10032005\I508658C.raw**

```
63 ctg cag ccc atc ctg gac agc ttc ccc aaa gat gtg gac ctc agc cag      460
64 Leu Gln Pro Ile Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln
65         95                  100                 105
67 ccc cgg aag ggg agg aag ccc ccg gcc gtc ccc aag gct ttg gta ccg      508
68 Pro Arg Lys Gly Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro
69     110                 115                 120
71 cca ccc aga ctc ccc acc aag agg aag gcc tca gaa gag gct cga gct      556
72 Pro Pro Arg Leu Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala
73 125                 130                 135                 140
75 gcc gcg cca gca gcc ctg act cca agg ggc acc gcc agc cca ggc tct      604
76 Ala Ala Pro Ala Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser
77                 145                 150                 155
79 caa ctg aag gcc aag ccc ccc aag aag ccg gag agc agc gca gag cag      652
80 Gln Leu Lys Ala Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln
81             160                 165                 170
83 cag cgc ctt cca ctc ggg aac ggg att cag acc atg tca gct tca gtc      700
84 Gln Arg Leu Pro Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val
85         175                 180                 185
87 cag aga gct gtg gcc atg tcc tcc ggg gac gtc ccg gga gcc cga ggg      748
88 Gln Arg Ala Val Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg Gly
89     190                 195                 200
91 gcc gtg gag ggg atc ctc atc cag cag gtg ttt gag tca ggc ggc tcc      796
92 Ala Val Glu Gly Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser
93 205                 210                 215                 220
95 aag aag tgc atc cag gtt ggc ggg gag ttc tac act ccc agc aag ttc      844
96 Lys Lys Cys Ile Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys Phe
97                 225                 230                 235
99 gaa gac tcc ggc agt ggg aag aac aag gcc cgc agc agc agt ggc ccg      892
100 Glu Asp Ser Gly Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro
101             240                 245                 250
103 aag cct ctg gtt cga gcc aag gga gcc cag ggc gct gcc ccc ggt gga      940
104 Lys Pro Leu Val Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly
105         255                 260                 265
107 ggt gag gct agg ctg ggc cag cag ggc agc gtt ccc gcc cct ctg gcc      988
108 Gly Glu Ala Arg Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala
109     270                 275                 280
111 ctc ccc agt gac ccc cag ctc cac cag aag aat gag gac gag tgt gcc     1036
112 Leu Pro Ser Asp Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala
113 285                 290                 295                 300
115 gtg tgt cgg gac ggc ggg gag ctc atc tgc tgt gac ggc tgc cct cgg     1084
116 Val Cys Arg Asp Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg
117                 305                 310                 315
119 gcc ttc cac ctg gcc tgc ctg tcc cct ccg ctc cgg gag atc ccc agt     1132
120 Ala Phe His Leu Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser
121             320                 325                 330
123 ggg acc tgg agg tgc tcc agc tgc ctg cag gca aca gtc cag gag gtg     1180
124 Gly Thr Trp Arg Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val
125         335                 340                 345
127 cag ccc cgg gca gag gag ccc cgg ccc cag gag cca ccc gtg gag acc     1228
```

RAW SEQUENCE LISTING DATE: 10/03/2005
PATENT APPLICATION: **US/09/508,658C** TIME: 13:29:19

Input Set : **A:\u012653-9.txt**
Output Set: **N:\CRF4\10032005\I508658C.raw**

```
128 Gln Pro Arg Ala Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr
129     350                 355                 360
131 ccg ctc ccc ccg ggg ctt agg tcg gcg gga gag gag gta aga ggt cca     1276
132 Pro Leu Pro Pro Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro
133 365                 370                 375                 380
135 cct ggg gaa ccc cta gcc ggc atg gac acg act ctt gtc tac aag cac     1324
136 Pro Gly Glu Pro Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His
137                 385                 390                 395
139 ctg ccg gct ccg cct tct gca gcc ccg ctg cca ggg ctg gac tcc tcg     1372
140 Leu Pro Ala Pro Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser
141             400                 405                 410
143 gcc ctg cac ccc cta ctg tgt gtg ggt cct gag ggt cag cag aac ctg     1420
144 Ala Leu His Pro Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu
145         415                 420                 425
147 gct cct ggt gcg cgt tgc ggg gtg tgc gga gat ggt acg gac gtg ctg     1468
148 Ala Pro Gly Ala Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu
149     430                 435                 440
151 cgg tgt act cac tgc gcc gct gcc ttc cac tgg cgc tgc cac ttc cca     1516
152 Arg Cys Thr His Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro
153 445                 450                 455                 460
155 gcc ggc acc tcc cgg ccc ggg acg ggc ctg cgc tgc aga tcc tgc tca     1564
156 Ala Gly Thr Ser Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser
157                 465                 470                 475
159 gga gac gtg acc cca gcc cct gtg gag ggg gtg ctg gcc ccc agc ccc     1612
160 Gly Asp Val Thr Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro
161             480                 485                 490
163 gcc cgc ctg gcc cct ggg cct gcc aag gat gac act gcc agt cac gag     1660
164 Ala Arg Leu Ala Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu
165         495                 500                 505
167 ccc gct ctg cac agg gat gac ctg gag tcc ctt ctg agc gag cac acc     1708
168 Pro Ala Leu His Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr
169     510                 515                 520
171 ttc gat ggc atc ctg cag tgg gcc atc cag agc atg gcc cgt ccg gcg     1756
172 Phe Asp Gly Ile Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala
173 525                 530                 535                 540
175 gcc ccc ttc ccc tcc tgaccccaga tggccgggac atgcagctct gatgagagag     1811
176 Ala Pro Phe Pro Ser
177                 545
179 tgctgagaag gacacctcct tcctcagtcc tggaagccgg ccggctggga tcaagaaggg   1871
181 gacagcgcca cctcttgtca gtgctcggct gtaaacagct ctgtgtttct ggggacacca   1931
183 gccatcatgt gcctggaaat taaaccctgc cccacttctc tactctggaa gtccccggga   1991
185 gcctctcctt gcctggtgac ctactaaaaa tataaaaatt agctg                   2036
188 <210> SEQ ID NO: 2
189 <211> LENGTH: 545
190 <212> TYPE: PRT
191 <213> ORGANISM: HOMO SAPIENS
193 <400> SEQUENCE: 2
195 Met Ala Thr Asp Ala Ala Leu Arg Arg Leu Leu Arg Leu His Arg Thr
196 1               5                   10                  15
```

RAW SEQUENCE LISTING DATE: 10/03/2005
PATENT APPLICATION: US/09/508,658C TIME: 13:29:19

Input Set : A:\u012653-9.txt
Output Set: N:\CRF4\10032005\I508658C.raw

```
199 Glu Ile Ala Val Ala Val Asp Ser Ala Phe Pro Leu Leu His Ala Leu
200             20                  25                  30
203 Ala Asp His Asp Val Val Pro Glu Asp Lys Phe Gln Glu Thr Leu His
204         35                  40                  45
207 Leu Lys Glu Lys Glu Gly Cys Pro Gln Ala Phe His Ala Leu Leu Ser
208     50                  55                  60
211 Trp Leu Leu Thr Gln Asp Ser Thr Ala Ile Leu Asp Phe Trp Arg Val
212 65                  70                  75                  80
215 Leu Phe Lys Asp Tyr Asn Leu Glu Arg Tyr Gly Arg Leu Gln Pro Ile
216                 85                  90                  95
219 Leu Asp Ser Phe Pro Lys Asp Val Asp Leu Ser Gln Pro Arg Lys Gly
220             100                 105                 110
223 Arg Lys Pro Pro Ala Val Pro Lys Ala Leu Val Pro Pro Pro Arg Leu
224         115                 120                 125
227 Pro Thr Lys Arg Lys Ala Ser Glu Glu Ala Arg Ala Ala Ala Pro Ala
228     130                 135                 140
231 Ala Leu Thr Pro Arg Gly Thr Ala Ser Pro Gly Ser Gln Leu Lys Ala
232 145                 150                 155                 160
235 Lys Pro Pro Lys Lys Pro Glu Ser Ser Ala Glu Gln Gln Arg Leu Pro
236                 165                 170                 175
239 Leu Gly Asn Gly Ile Gln Thr Met Ser Ala Ser Val Gln Arg Ala Val
240             180                 185                 190
243 Ala Met Ser Ser Gly Asp Val Pro Gly Ala Arg Gly Ala Val Glu Gly
244         195                 200                 205
247 Ile Leu Ile Gln Gln Val Phe Glu Ser Gly Gly Ser Lys Lys Cys Ile
248     210                 215                 220
251 Gln Val Gly Gly Glu Phe Tyr Thr Pro Ser Lys Phe Glu Asp Ser Gly
252 225                 230                 235                 240
255 Ser Gly Lys Asn Lys Ala Arg Ser Ser Ser Gly Pro Lys Pro Leu Val
256                 245                 250                 255
259 Arg Ala Lys Gly Ala Gln Gly Ala Ala Pro Gly Gly Gly Glu Ala Arg
260             260                 265                 270
263 Leu Gly Gln Gln Gly Ser Val Pro Ala Pro Leu Ala Leu Pro Ser Asp
264         275                 280                 285
267 Pro Gln Leu His Gln Lys Asn Glu Asp Glu Cys Ala Val Cys Arg Asp
268     290                 295                 300
271 Gly Gly Glu Leu Ile Cys Cys Asp Gly Cys Pro Arg Ala Phe His Leu
272 305                 310                 315                 320
275 Ala Cys Leu Ser Pro Pro Leu Arg Glu Ile Pro Ser Gly Thr Trp Arg
276                 325                 330                 335
279 Cys Ser Ser Cys Leu Gln Ala Thr Val Gln Glu Val Gln Pro Arg Ala
280             340                 345                 350
283 Glu Glu Pro Arg Pro Gln Glu Pro Pro Val Glu Thr Pro Leu Pro Pro
284         355                 360                 365
287 Gly Leu Arg Ser Ala Gly Glu Glu Val Arg Gly Pro Pro Gly Glu Pro
288     370                 375                 380
291 Leu Ala Gly Met Asp Thr Thr Leu Val Tyr Lys His Leu Pro Ala Pro
292 385                 390                 395                 400
295 Pro Ser Ala Ala Pro Leu Pro Gly Leu Asp Ser Ser Ala Leu His Pro
```

RAW SEQUENCE LISTING DATE: 10/03/2005
PATENT APPLICATION: US/09/508,658C TIME: 13:29:19

Input Set : A:\u012653-9.txt
Output Set: N:\CRF4\10032005\I508658C.raw

```
296                 405                 410                 415
299 Leu Leu Cys Val Gly Pro Glu Gly Gln Gln Asn Leu Ala Pro Gly Ala
300             420                 425                 430
303 Arg Cys Gly Val Cys Gly Asp Gly Thr Asp Val Leu Arg Cys Thr His
304         435                 440                 445
307 Cys Ala Ala Ala Phe His Trp Arg Cys His Phe Pro Ala Gly Thr Ser
308     450                 455                 460
311 Arg Pro Gly Thr Gly Leu Arg Cys Arg Ser Cys Ser Gly Asp Val Thr
312 465                 470                 475                 480
315 Pro Ala Pro Val Glu Gly Val Leu Ala Pro Ser Pro Ala Arg Leu Ala
316                 485                 490                 495
319 Pro Gly Pro Ala Lys Asp Asp Thr Ala Ser His Glu Pro Ala Leu His
320             500                 505                 510
323 Arg Asp Asp Leu Glu Ser Leu Leu Ser Glu His Thr Phe Asp Gly Ile
324         515                 520                 525
327 Leu Gln Trp Ala Ile Gln Ser Met Ala Arg Pro Ala Ala Pro Phe Pro
328     530                 535                 540
331 Ser
332 545
335 <210> SEQ ID NO: 3
336 <211> LENGTH: 1545
337 <212> TYPE: DNA
338 <213> ORGANISM: HOMO SAPIENS
341 <220> FEATURE:
342 <221> NAME/KEY: CDS
343 <222> LOCATION: (237)..(1280)
345 <400> SEQUENCE: 3
346 agagaaagtg aggtcttctc aggctcttaa gagcatggcg tttggtccag gctgtacccg     60
348 ctgctctcag ctgggcccgt gggtgggccg ggcgcccctg ctatagccag gaggtcaagg    120
350 atccactggg aatgccatgc tcatctttcg tccccagcat ggtttcttaa tggggtagaa    180
352 gcaggtcggg agagacctcc ctgggcctgg ccccactgcc ctgtgaggaa gggttc atg    239
353                                                                 Met
354                                                                 1
356 tgg ttg gtg tac agt tcc ggg gcc cct gga acg cag cag cct gca aga      287
357 Trp Leu Val Tyr Ser Ser Gly Ala Pro Gly Thr Gln Gln Pro Ala Arg
358             5                   10                  15
360 aac cgg gtt ttc ttc cca ata ggg atg gcc ccg ggg ggt gtc tgt tcg      335
361 Asn Arg Val Phe Phe Pro Ile Gly Met Ala Pro Gly Gly Val Cys Ser
362         20                  25                  30
364 aga cca gat gga tgg gga aca ggt ggt cag ggc aga att tca ggc cct      383
365 Arg Pro Asp Gly Trp Gly Thr Gly Gly Gln Gly Arg Ile Ser Gly Pro
366     35                  40                  45
368 ggc agc atg gga gca ggg cag aga ctg ggg agt tca ggt acc cag aga      431
369 Gly Ser Met Gly Ala Gly Gln Arg Leu Gly Ser Ser Gly Thr Gln Arg
370 50                  55                  60                  65
372 tgc tgc tgg ggg agc tgt ttt ggg aag gag gtg gct ctc agg agg gtg      479
373 Cys Cys Trp Gly Ser Cys Phe Gly Lys Glu Val Ala Leu Arg Arg Val
374                 70                  75                  80
376 ctg cac ccc agc cca gtc tgc atg ggc gtc tct tgc ctg tgc cag aag      527
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 10/03/2005
PATENT APPLICATION: **US/09/508,658C** TIME: 13:29:20

Input Set : **A:\u012653-9.txt**
Output Set: **N:\CRF4\10032005\I508658C.raw**

Invalid <213> Response:
Use of "Artificial" only as "<213> Organism" response is incomplete, per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:7,8,9,10,11,12,13,14,15,16,17,18,19,20,21,22,23,24,25,26,27,28,29,30,31
Seq#:32,33,34,35

VERIFICATION SUMMARY DATE: 10/03/2005
PATENT APPLICATION: **US/09/508,658C** TIME: 13:29:20

Input Set : **A:\u012653-9.txt**
Output Set: **N:\CRF4\10032005\I508658C.raw**